



## Prediction of speech masking release for fluctuating interferers based on the envelope power signal-to-noise ratio

Jørgensen, Søren; Dau, Torsten

*Published in:*  
Journal of the Acoustical Society of America

*Link to article, DOI:*  
[10.1121/1.4708508](https://doi.org/10.1121/1.4708508)

*Publication date:*  
2012

[Link back to DTU Orbit](#)

*Citation (APA):*  
Jørgensen, S., & Dau, T. (2012). Prediction of speech masking release for fluctuating interferers based on the envelope power signal-to-noise ratio. *Journal of the Acoustical Society of America*, 131, 3341-3341. <https://doi.org/10.1121/1.4708508>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**Title:**

Prediction of speech masking release for fluctuating interferers based on the envelope power signal-to-noise ratio.

**Authors:** Søren Jørgensen and Torsten Dau.

**Abstract:**

The speech-based envelope power spectrum model (sEPSM) presented by Jørgensen and Dau [(2011). *J. Acoust. Soc. Am.* **130**, 1475-1487] estimates the envelope signal-to-noise ratio ( $\text{SNR}_{\text{env}}$ ) after modulation-frequency selective processing, which accurately predicts the speech intelligibility for normal-hearing listeners in conditions with additive stationary noise, reverberation, and nonlinear processing with spectral subtraction. The latter condition represents a case in which the standardized speech intelligibility index and speech transmission index fail. However, the sEPSM is limited to conditions with stationary interferers due to the long-term estimation of the envelope power and cannot account for the well known phenomenon of speech masking release. Here, a short-term version of the sEPSM is presented, estimating the envelope SNR in 10-ms time frames. Predictions obtained with the short-term sEPSM are compared to data from Kjems *et al.* [(2009). *J. Acoust. Soc. Am.* **126** (3), 1415-1426] where speech is mixed with four different interferers, including speech-shaped noise, bottle noise, car noise, and a highly non-stationary cafe noise. The model accounts well for the differences in intelligibility observed for the stationary and non-stationary interferers, demonstrating further that the envelope SNR is crucial for speech comprehension.